WILEY

RESEARCH ARTICLE

Intermittent theta burst stimulation over right somatosensory larynx cortex enhances vocal pitch-regulation in nonsingers

Sebastian Finkel[1] | Ralf Veit[1] | Martin Lotze[2] | Anders Friberg[3] | Peter Vuust[4] |
Surjo Soekadar[5,6] | Niels Birbaumer[1,7] | Boris Kleber[1,4]

[1]Institute of Medical Psychology and Behavioral Neurobiology, Eberhard Karls University Tübingen, Tübingen, Germany

[2]Functional Imaging Unit; Center for Diagnostic Radiology and Neuroradiology, University of Greifswald, Greifswald, Germany

[3]Department of Speech, Music and Hearing, KTH Royal Institute of Technology, Stockholm, Sweden

[4]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[5]Department of Psychiatry and Psychotherapy and Neuroscience Research Center (NWFZ), Charité Campus Mitte (CCM), Berlin, Germany

[6]Department of Psychiatry and Psychotherapy, University Hospital of Tübingen, Tübingen, Germany

[7]Wyss Center for Bio and Neuroengineering, Geneva, Switzerland

Correspondence
Boris Kleber, Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, Nørrebrogade 44, 10 G 8000 Aarhus, Denmark.
Email: boris.kleber@clin.au.dk

Funding information
European Union, Grant/Award Numbers: Project WAY (Wearable interfaces for hAnd function recoverY, Grant Number 288551), Project BRAIN TRAIN (Grant Number 602186); European Research Council; Brain & Behavior Research Foundation; EU; Deutsche Forschungsgemeinschaft, Grant/Award Number: KL 2341/1

Abstract

While the significance of auditory cortical regions for the development and maintenance of speech motor coordination is well established, the contribution of somatosensory brain areas to learned vocalizations such as singing is less well understood. To address these mechanisms, we applied intermittent theta burst stimulation (iTBS), a facilitatory repetitive transcranial magnetic stimulation (rTMS) protocol, over right somatosensory larynx cortex (S1) and a nonvocal dorsal S1 control area in participants without singing experience. A pitch-matching singing task was performed before and after iTBS to assess corresponding effects on vocal pitch regulation. When participants could monitor auditory feedback from their own voice during singing (Experiment I), no difference in pitch-matching performance was found between iTBS sessions. However, when auditory feedback was masked with noise (Experiment II), only larynx-S1 iTBS enhanced pitch accuracy (50–250 ms after sound onset) and pitch stability (>250 ms after sound onset until the end). Results indicate that somatosensory feedback plays a dominant role in vocal pitch regulation when acoustic feedback is masked. The acoustic changes moreover suggest that right larynx-S1 stimulation affected the preparation and involuntary regulation of vocal pitch accuracy, and that kinesthetic-proprioceptive processes play a role in the voluntary control of pitch stability in nonsingers. Together, these data provide evidence for a causal involvement of right larynx-S1 in vocal pitch regulation during singing.

KEYWORDS

predictive coding, sensorimotor, singing, TMS, vocal production

1 | INTRODUCTION

When people listen to their own voice through headphones while they speak or sing, they involuntarily adjust their vocal output when the pitch they hear is unexpectedly shifted upward or downward. This pitch-shift reflex has been described as an important indicator for a close relationship between the auditory feedback and the laryngeal motor systems (Burnett, Freedland, Larson, & Hain, 1998; Hain et al., 2000; Larson, Altman, Liu, & Hain, 2008; Leydon, Bauer, & Larson, 2003; Liu, Behroozmand, Bove, & Larson, 2011). Acoustic self-monitoring plays not only a fundamental role in vocal motor control to match specific acoustic targets (Hickok, Houde, & Rong, 2011) but can also be enhanced through training in skilled musicians (Parkinson et al., 2014). However, although vocal production also utilizes somatosensory feedback arising as sensations of touch, stretch, vibration, and position during respiratory, laryngeal, and articulatory activity (Andreatta, Mann, Poletto, & Ludlow, 2002; Berry, Montequin, Chan, Titze, & Hoffman, 2003; Gozaine & Clark, 2005; Hammer & Barlow, 2010; Jurgens, 2002;

Sundberg, Iwarsson, & Billstrom, 1995; Wyke, 1974a; Wyke, 1974b), the contribution of somatosensory processes to learned vocalizations is less well understood (Breshears, Molinaro, & Chang, 2015; Hammer & Barlow, 2010; Kleber, Zeitouni, Friberg, & Zatorre, 2013; Lametti, Nasir, & Ostry, 2012; Sundberg et al., 1995).

Current models of speech motor control describe and formalize the sensory-motor transformations that occur during speech production (Parrell, Lammert, Ciccarelli, & Quatieri, 2017). The directions into velocities of articulators (DIVA) model (Guenther, 2006; Tourville & Guenther, 2011) combines predictive processes with auditory and somatosensory feedback control loops. Its basic feature is a speech sound map located in ventral premotor cortex/posterior inferior frontal gyrus (vPMC/pIFG) that contains articulatory motor trajectories as input to a predictive controller, and corresponding sensory expectations as input to feedback controllers. The feedback controllers compare the expectations with incoming feedback and generate an error signal in case of mismatch. During early stages of vocal learning, auditory feedback is crucial to establish associations between motor commands and their acoustic consequences. With experience, however, feedforward mechanisms are tuned through sensory feedback that compute motor commands from predicted sensory consequences with sufficient accuracy to produce normal speech. At this point, auditory feedback is mainly used to maintain or adjust the feedforward model (see also Guenther, 2016). Similarly, state feedback control models suggest that rapid online speech control is achieved via internally represented predictions of sensory consequences of intended vocalizations, whereas the actual feedback during overt production is used to train and update the internal forward model in case of error detection (Hickok et al., 2011; Houde & Nagarajan, 2011). Neuroimaging studies indicate that the comparison between predicted and actual auditory feedback involves primary as well as higher-order auditory regions in posterior superior temporal gyrus (STG) and ventral supramarginal gyrus (SMG) that interact with vPMC (including Broca's area) to generate state corrections, whereas the parietal lobe, cerebellum, and basal ganglia contribute to the programming and maintenance of feedforward control (Behroozmand et al., 2016; Behroozmand et al., 2018; Behroozmand & Sangtian, 2018; Guenther & Vladusich, 2012; Hickok, 2017; Houde & Chang, 2015; Parkinson et al., 2012; Shum, Shiller, Baum, & Gracco, 2011; Tourville, Reilly, & Guenther, 2008; Zarate & Zatorre, 2008).

The neural mechanisms underlying the contribution of somatosensory feedback to voluntary vocal motor control are less well understood (Mor, Simonyan, & Blitzer, 2018). Research in cats, humans, and nonhuman primates suggests that motor nuclei in the brainstem (Kuypers, 1958a; Kuypers, 1958b; Kuypers, 1958c; Yoshida, Mitsumasu, Hirano, & Kanaseki, 1985) use dynamic feedback from laryngeal mechanoreceptors to stabilize vocal motor patterns by coordinating respiratory and phonatory muscle groups based on intrinsic laryngeal reflex mechanisms, whereas somatosensory feedback also provides information about the spatial position of the respiratory and vocal tracts (Ambalavanar, Tanaka, Selbie, & Ludlow, 2004; Andreatta et al., 2002; Borich, Brodie, Gray, Ionta, & Boyd, 2015; Gozaine & Clark, 2005; Ludlow, 2005; Nasir & Ostry, 2006; Simonyan & Horwitz, 2011; Smotherman, 2007; Wyke, 1974a). The DIVA model predicts that self-monitoring of correctly produced sounds generates somatosensory targets throughout speech motor development, representing

proprioceptive and tactile sensations of the simultaneously generated sound (Guenther & Vladusich, 2012; Ito, Coppola, & Ostry, 2016; Skipper, Devlin, & Lametti, 2017; Tremblay, Shiller, & Ostry, 2003). This implies that auditory and somatosensory feedback control subsystems become tightly coupled (Ito & Ostry, 2012; Ito, Tiede, & Ostry, 2009). In keeping with this notion, studies have shown that the magnitude of vocal responses to pitch perturbations may be constrained by the integration of somatosensory feedback (Katseff, Houde, & Johnson, 2012; Larson et al., 2008). Moreover, individuals also generate compensatory responses to unexpected somatosensory (e.g., jaw) perturbations (Lametti et al., 2012; Nasir & Ostry, 2006; Nasir & Ostry, 2008). The underlying neural computations activate ventral primary somatosensory cortex (S1) and adjacent SMG interacting with vPMC (Golfinopoulos et al., 2011). Electrophysiological and neurostimulation studies have moreover demonstrated that the area in and around the SMG, which receives projections from various primary sensory areas, is involved in sensorimotor adaptation during speech production (Behroozmand & Sangtian, 2018; Shum et al., 2011).

In order to assess experience-dependent effects and their neural correlates, pitch-matching tasks with singers and nonsingers provide an interesting model (Jones & Keough, 2008; Kleber et al., 2013; Kleber, Friberg, Zeitouni, & Zatorre, 2017; Zarate, 2013; Zarate, Wood, & Zatorre, 2010), as even the slightest deviation from musically defined target sounds will be perceived as error (Hutchins & Peretz, 2012; Natke, Donath, & Kalveram, 2003; Zatorre & Baum, 2012). To date, only few neuroimaging studies with trained singers have been performed. They found that singing experience can be associated with increased activation of ventral S1, inferior parietal lobe (IPL), and the cerebellum during normal singing (Kleber, Veit, Birbaumer, Gruzelier, & Lotze, 2010). The interaction between singing experience and auditory perturbation moreover leads to increased activation of IPL in singers together with ventral inferior frontal gyrus (vIFG, BA44) and supplementary motor area (Kleber et al., 2017). This led to the hypothesis that singing training may improve the accuracy of somatosensory feedback for regulating singing voice output (Kleber et al., 2017; Mürbe, Pabst, Hofmann, & Sundberg, 2004). Singers may thus possess more accurate somatosensory targets and more robust feedforward representations (Jones & Keough, 2008; Kleber et al., 2013; Villacorta, Perkell, & Guenther, 2007), which explains their reduced sensitivity to auditory feedback perturbations (Kleber et al., 2017; Zarate & Zatorre, 2008).

Against this background, the current study set out to explore if enhanced somatosensory feedback processing can improve vocal pitch control in nonsingers. To achieve this goal, we used intermitted theta burst stimulation (iTBS), a form of repetitive transcranial magnetic stimulation (rTMS), to facilitate cortical processing in right ventral larynx S1 (Huang, Edwards, Rounis, Bhatia, & Rothwell, 2005; Jones et al., 2016; Katayama & Rothwell, 2007; Morley, Vickery, Stuart, & Turman, 2007). The right hemisphere was chosen based on studies indicating that singing compared to speaking activates a larger bi-hemispheric network, reflecting additional right lateralized activation and gray matter volume related to vocal production within a musical context (Kleber et al., 2016; Özdemir, Norton, & Schlaug, 2006; Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000).

Feedback error detection and sensorimotor adaptation are moreover processed bilaterally (Golfinopoulos et al., 2011; Parkinson et al., 2012; Tourville et al., 2008), with right parietal and premotor regions contributing to the coordination of compensatory speech motor responses together with left posterior STG (Kort, Cuesta, Houde, & Nagarajan, 2016). Enhanced right-hemisphere activation has also been associated with superior vocal pitch processing in musicians with enhanced pitch identification mechanisms (Behroozmand, Ibrahim, Korzyukov, Robin, & Larson, 2014; Behroozmand, Ibrahim, Korzyukov, Robin, & Larson, 2015; Parkinson et al., 2014). A previous rTMS study demonstrated that the SMG is involved in adaptive vocal responses to feedback alterations (Shum et al., 2011). However, the current study aimed at enhancing vocalization-related somatosensory perception and therefore targeted directly the somatotopically organized larynx-S1 (Guenther & Vladusich, 2012). Previous studies that used facilitating iTBS protocols over S1 reported enhanced tactile perception (Ragert, Franzkowiak, Schwenkreis, Tegenthoff, & Dinse, 2008), whereas inhibitory continuous TBS impairs both tactile spatial (Rai, Premji, Tommerdahl, & Nelson, 2012) and temporal acuity (Lee et al., 2013; Rai et al., 2012). S1 is moreover heavily connected to both the inferior parietal cortex (Borich et al., 2015) and somatotopically associated with M1, integrating motor commands with proprioceptive and tactile information even before vocalization (Bouchard, Mesgarani, Johnson, & Chang, 2013).

To assess the effects of neuronavigated iTBS over right ventral-S1 on singing, nonsingers performed a pitch-matching singing paradigm before and after stimulation. The effect was compared to stimulation of dorsal S1, which is not part of the singing network. In Experiment I, participants monitored auditory feedback from their own voice during singing. We expected that larynx-S1 but not the control stimulation would improve pitch-matching performance in nonsingers, indicating a role of somatosensory feedback in vocal pitch regulation. Based on evidence indicating that nonsingers rely more heavily on auditory feedback pitch regulation in a musical context (Jones & Keough, 2008; Kleber et al., 2013; Kleber et al., 2017; Zarate & Zatorre, 2008), a second experiment was performed in which auditory self-monitoring was masked with noise. We hypothesized that that the lack of auditory feedback may trigger the monitoring system to attend more to somatosensory feedback. Therefore, we expected greater effects of iTBS over right ventral S1 on pitch-matching performance in Experiment II.

2 | MATERIAL AND METHODS

2.1 | Design and participants

Two related experiments were performed that differed with respect to the presence or absence of pink noise during vocal pitch-matching performance (Figure 1), which aimed to mask auditory-feedback from the own voice. Participants were randomly recruited from the University of Tübingen and matched for sex in both experiments. All subjects reported no history of neurological, speech or hearing disorder. TMS specific inclusion criteria were applied according to Wassermann (Wassermann, 1998). Only participants without prior formal music

training (singing or playing an instrument) were included in this study. Occasional musical activities (e.g., choirs or informal bands) served as additional exclusion criteria. In Experiment I, the effects of iTBS on pitch matching accuracy was tested while participants monitored auditory feedback to control their vocal output. In line with a two-stage experimental design (Simon, 1989), recruitment in Experiment I was stopped after nine participants (four females; mean age = 24.9 years; age range = 20–33 years) due to a low effect size. As no participant showed significant effects of iTBS on pitch-matching performance, the inclusion of more participants was unlikely to change the outcome. Experiment II included 14 participants (seven females) with a mean age of 27.3 years (range: 22–35 years). Four subjects who previously participated in Experiments I, also volunteered to participate in Experiment II (one female, three male). Pitch-matching accuracy can be considered relatively stable (Zarate, Delhommeau, Wood, & Zatorre, 2010). Therefore, we expected no heterogeneity due to potential test–retest effects in the four participants. In Experiment II, auditory feedback from the own voice was masked with pink noise during singing to reveal effects of iTBS on pitch performance. In addition, participants also sang with normal auditory feedback (only prior to stimulation) in Experiment II. This was done to validate the predicted effects of masking noise on pitch performance (i.e., reduced accuracy). The ethics committee of the Medical Faculty of the University of Tübingen approved this study and written consent was obtained prior to participation.

2.2 | Pitch-discrimination accuracy

A two-tone forced choice pitch-discrimination test was initially performed to rule out that poor pitch perception could affect pitch-matching performance. Pure tones with a duration of 250 ms, separated by 600 ms gaps, and an initial frequency difference Δf = 7% were presented via Bose Quiet-15 headphones. The lower standard tone was fixed at 500 Hz while the order of both tones was randomized. During the test, Δf was adaptively changed using a two-down one-up rule, which tracks 70.7%-correct thresholds on the psychometric function (Levitt, 1971). Δf was decreased after two consecutive correct responses and increased after one incorrect response by the factor β. Initially, β was set to 2 but changed to 1.25 after the second reversal. The test was terminated after 15 reversals and the final threshold was calculated as the geometric mean of Δf across the last eight reversals. A pitch discrimination threshold exceeding a 2% frequency difference, corresponding to a 35 Cents interval (100 Cents = 1 semitone) served as exclusion criteria for the participant. All participants passed this test, indicating that precipitants' perceptual thresholds were at least 65% smaller than the smallest pitch interval required for pitch-matching in this study (1 semitone). Typical frequency discrimination thresholds are 0.86% in healthy nonmusicians and 0.13% in trained musicians (Micheyl, Delhommeau, Perrot, & Oxenham, 2006).

2.3 | Pitch-matching performance

During pitch-matching, participants first listened to musical intervals (one pair of tones per trial) and subsequently intended to match the fundamental frequencies of these tones by singing the syllable/da/.
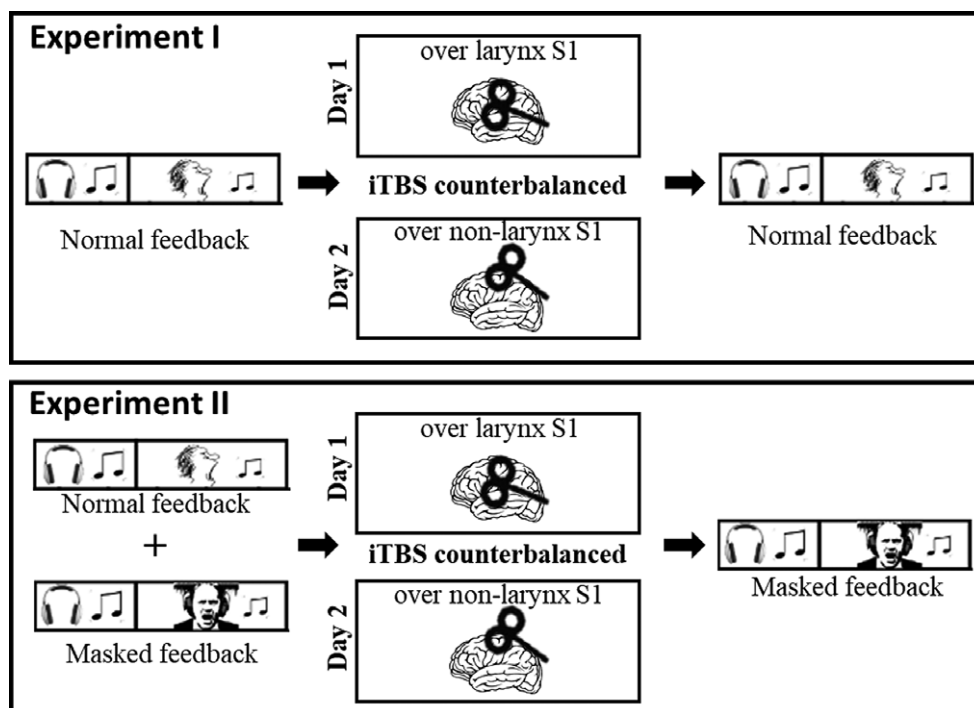
FIGURE 1 The experimental procedures for Experiments I and II are depicted. Neuronavigated iTBS was applied in two pseudorandomized sessions over right larynx-S1 and over a nonvocal dorsal-S1 control area (representing the lower limb) with a minimum of 48 hr washout time between them. Effects on pitch-matching performance were assessed before and after stimulation. In Experiment I, participants could use auditory feedback from their own voice to monitor vocal pitch. In experiment II, loud pink noise masked auditory self-monitoring. In addition, pitch-matching was also performed with normal auditory feedback (only prior to iTBS)

A voiced syllable was chosen to prevent participants from performing a glissando between tones. Each tone of the interval was presented via Bose Quiet15 headphones with a duration of 900 ms, separated by a 200 ms gap, approximating the timbre of the vowel/a/ of a human voice (Hutchins & Peretz, 2011). The first tone was fixed (311.11 Hz = $D^{\#4}$ for females and 155.56 Hz = $D^{\#3}$ for males) while the second tone differed by ±1–9 semitones. Five practice trials preceded the experiment.

In Experiment I, 105 pitch-matching trials were performed, once before and once again after iTBS. During singing, participants could use auditory feedback from their own voice to monitor vocal pitch. In Experiment II, loud pink noise was played via Bose Quiet15 headphones while participants intended to match the musical intervals to reduce auditory self-monitoring during pitch-matching. This was done once before and once again after iTBS. In order to validate the intended effects of masking noise on pitch performance (i.e., reduced accuracy), participants also sang without masking (prior to iTBS) in Experiment II. The number of trials was reduced to 60 in Experiment II for reasons of design economy, as the lower number provided sufficient statistical power for detecting changes in pitch-matching performance (Hutchins & Peretz, 2012).

Pink noise is typically used in pitch-shifting paradigms to reduce air- and bone-conduction feedback from the own voice (Bauer & Larson, 2003). Neuroimaging studies demonstrated that the perceived loudness of pink noise is linearly related to neural activation levels in auditory cortex (Rohl, Kollmeier, & Uppenkamp, 2011) and successfully prevents vocalization-related cortical potentials (Masaki, Tanaka, Takasawa, & Yamazaki, 2001). Noise levels were individually adjusted to the highest tolerable level (below pain thresholds) to limit the perception of residual auditory feedback. Participants were furthermore instructed to sing with low volume to maximize the masking effect.

Vocal responses were captured with a head mounted microphone (C 477 WR L/p, AKG, Austria) connected to an M-Audio firewire 410 external sound card (inMusic, Ratingen, Germany) to ensure equidistant microphone positioning and stored in wav-format for later off-line analyses. Stimuli presentation and voice recordings were performed within MAX/MSP software ('74 Cycling).

2.4 | Intermittent theta burst stimulation

We applied iTBS with a Magstim Rapid[2] stimulator attached to a 70 mm biphasic figure-of-eight coil (The Magstim Company Limited, Withland, UK) in combination with a Localite stereotactic neuronavigation system (Localite GmbH, Sankt Augustin, Germany). T1-weighted whole-brain anatomical images were used for anatomically guided iTBS application with a coil orientation anterior to posterior and perpendicular to the postcentral gyrus for S1 larynx stimulation (Jacobs, Premji, & Nelson, 2012; Pfannmoller, Schweizer, & Lotze, 2016; Roux, Djidjeli, & Durand, 2018; Thielscher, Opitz, & Windhoff, 2011). Based on previous findings (Thielscher et al., 2011), the coil was tilted by 45° for dorsal-iTBS control stimulation (lower limb representation) to further reduce the effect on the cortex and create a sham like stimulation with similar skin sensation. We used an fMRI localizer task to identify participants' individual larynx representation in right ventral S1 (for details see Section 2.5 below).
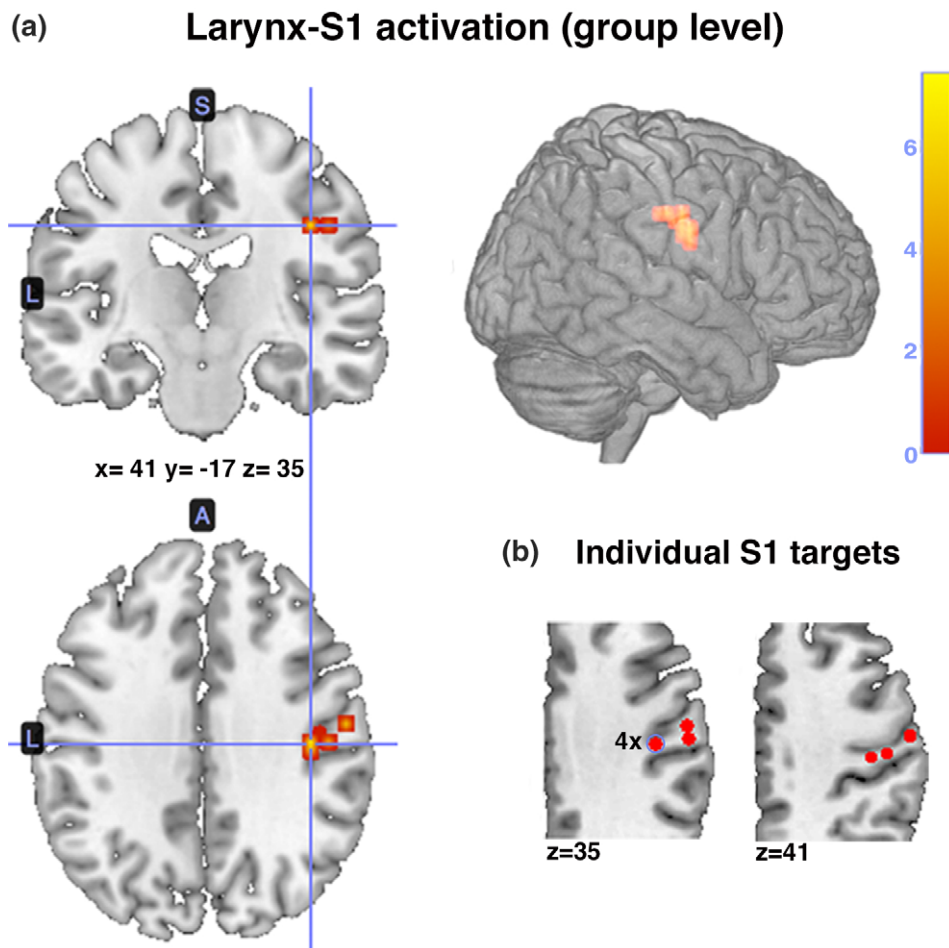
FIGURE 2 Results from the fMRI larynx-S1 localizer task in experiment 1. Participants sang pitch-glides on the steady vowel/a/ while keeping the articulators in a fixed position. The highest activated voxel in postcentral gyrus (small volume corrected; $p < 0.001$, >20vx) represented the target coordinates in right S1 (singing vs. rest). (a) Group activation ($N = 9$) in right S1 ($x = 41$, $y = -17$, $z = 35$ in MNI space) matched a previously reported larynx/phonation area, which was used for right larynx-S1 stimulation in Experiment II. (b) Individual S1 targets (highest activated voxel) in Experiment I are depicted. Coordinates were recomputed using Localite software to account for individual brain anatomy prior to neuronavigated iTBS, visually inspected and manually adjusted to spatially center the locus of stimulation on S1 [Color figure can be viewed at wileyonlinelibrary.com]

Theta burst stimulation (TBS) included three pulses of stimulation at 50 Hz repeated every 200 ms. In the intermittent stimulation pattern (iTBS), a 2 s train of TBS is repeated every 10 s for a total of 190 s (600 pulses; Huang et al., 2005). The order of stimulation (larynx and dorsal S1) was counterbalanced and performed on separate days to allow a minimum of 48 hr washout time between sessions. Stimulation output was set individually to 80% active motor threshold, determined prior to the initial session using motor evoked potentials (MEPs) of the right dorsal interossei (FDI) muscle. MEPs were recorded using 24 mm Ag/AgCl self-adhesive electrodes, with one electrode placed on the belly of the muscle and the other on muscle tendon. EMG signals were amplified, band-pass filtered (20–1,000 Hz), and sampled at 2 kHz. Criteria for active motor threshold followed a 5 out of 10 rule for pulse-evoked muscle potentials with a response larger than 100 μV (Bennemann, Freigang, Schroger, Rubsamen, & Richter, 2013). Participants were asked to keep their muscle tension constant at about 20% of maximum strength by pressing the thumb against the index finger. The mean motor threshold in Experiment I corresponded to 40% of the maximum stimulator output, which was used throughout Experiment II

(Kalla, Muggleton, Cowey, & Walsh, 2009; Leveque, Muggleton, Stewart, & Schon, 2013).

2.5 | Magnetic resonance imaging

T1-weighted anatomical images of the whole brain were acquired on a 3 T whole body Scanner (Siemens Magnetom Tim Trio, Erlangen, Germany) using a fast 3D gradient echo pulse sequence (MPRAGE; TR = 2,300 ms, TE = 2.98 ms, FA = 9°, matrix size = 248 × 256, FOV = 256 mm, voxel size = $1 \times 1 \times 1$ mm³).

Function echo planar imaging with a total of 72 whole-head volumes was used for the fMRI localizer task in Experiment 1 (echo time [TE]: 30 ms, repetition time [TR]: 10 s, including a 7 s sparse-temporal sampling delay in TR, actual volume acquisition time [TA]: 3 s, 40 transversal slices, 3.4 mm thickness; see Kleber et al., 2010, 2013; Kleber, Birbaumer, Veit, Trevorrow, & Lotze, 2007). This task required participants to sing pitch-glides on the steady vowel /a/. Individual coordinates of right ventral larynx-S1 were determined as the highest activated voxel in postcentral gyrus, representing the somatotopic maxima of healthy volunteers during singing compared with rest. As

the larynx-S1 group activation peak matched previously reported coordinates of a larynx/phonation area (Belyk & Brown, 2017; Brown et al., 2009; Brown, Ngan, & Liotti, 2008; Grabski et al., 2012; Kleber et al., 2010; Lotze, Seggewies, Erb, Grodd, & Birbaumer, 2000), these coordinates ($x = 41$, $y = -17$, $z = 35$ in MNI space) were used throughout Experiment II (Figure 2a). Figure 2b indicates the location of individual stimulation areas in Experiment 1.

For neuronavigated iTBS, coordinates were recomputed using Localite software to account for individual brain anatomy. Coordinates were visually inspected and manually adjusted to center its focal location with the aim of spatially limiting the locus of stimulation to S1 (Reichenbach, Thielscher, Peer, Bulthoff, & Bresciani, 2014).

2.6 | Acoustic voice analysis

A high-pass filter (Butterworth, order 4 or 8) was first set to the lowest expected sung frequency to filter out low frequency noise components. The remaining noise floor was estimated, onset and offset times of the two tones determined from the crossing points of the sound level envelope and the same envelope, low-pass filtered and a level offset included. These onset and offset positions estimated the start and end points of the voiced part of the sound. Audio signals were then divided into early (50–250 ms) and a late (250 ms-end) time components following sound-level onset detection by estimating the slope using a linear least-square polynomial fit. The early and late components are assumed to reflect subcortical and cortical control functions, respectively (Burnett, McCurdy, & Bright, 2008; Grell, Sundberg, Ternstrom, Ptok, & Altenmüller, 2009; Hain et al., 2000). The first 50 ms were omitted, as they typically exhibit large pitch variations.

Pitch-matching analysis was performed using a custom-made script within the CUEX performance analysis system run under Matlab (Friberg, Schoonderwaldt, & Juslin, 2005). Deviation from target-pitch was estimated as the median using the YIN algorithm

(de Cheveigne & Kawahara, 2002), limited to frequency range used in the experiment (C#3 to D5 for the females and F2 to B3 for the males), and expressed in cent for each tone (100 Cents = 1 semitone). Pitch stability was estimated using the interquartile range, where lower values indicate higher stability in tone production. Results were averaged across both tones (tone1 + tone2/2) using unsigned values.

2.7 | Statistical data analysis

Statistical analyses of pitch-performance before and after iTBS were performed with the statistical software package R (R Development Core Team 2014). Shapiro–Wilk tests indicated a non-normal data distribution. Therefore, nonparametric tests were employed throughout all analyses. Friedman's repeated measures ANOVA was used in to detect differences in pitch-matching performance before and after larynx-S1 iTBS and dorsal-S1 iTBS (representing the lower limb). In Experiment II, we first verified the predicted effects of masked auditory feedback on pitch-matching performance (i.e., decreased relative to normal auditory feedback) using Wilcoxon signed-rank tests. Thereafter, we used a robust version of Friedman's ANOVA (Skilling & Mack, 1981) to detect differences in pitch-matching performance with noise-masked auditory feedback before and after iTBS of larynx and dorsal S1 stimulation as in Experiment I.

In both experiments, one-sided Wilcoxon signed-rank tests were employed for post hoc pair-wise comparisons, corrected for multiple tests using false-discovery rate (FDR) with $p < 0.05$. The FDR method controls the expected proportion of false discoveries amongst the rejected hypotheses and is a less stringent condition than the family-wise error rate, so these methods are more powerful than the others (e.g., Bonferroni correction). We calculated and reported effect sizes for Wilcoxon signed-rank tests using this formula: $r = Z/\sqrt{N}$ (Rosenthal, 1994). Effect sizes cannot be calculated directly for Friedman tests. Finally, exploratory correlation analyses

Exp. I: Post *vs* pre iTBS differences in pitch performance (with normal hearing)
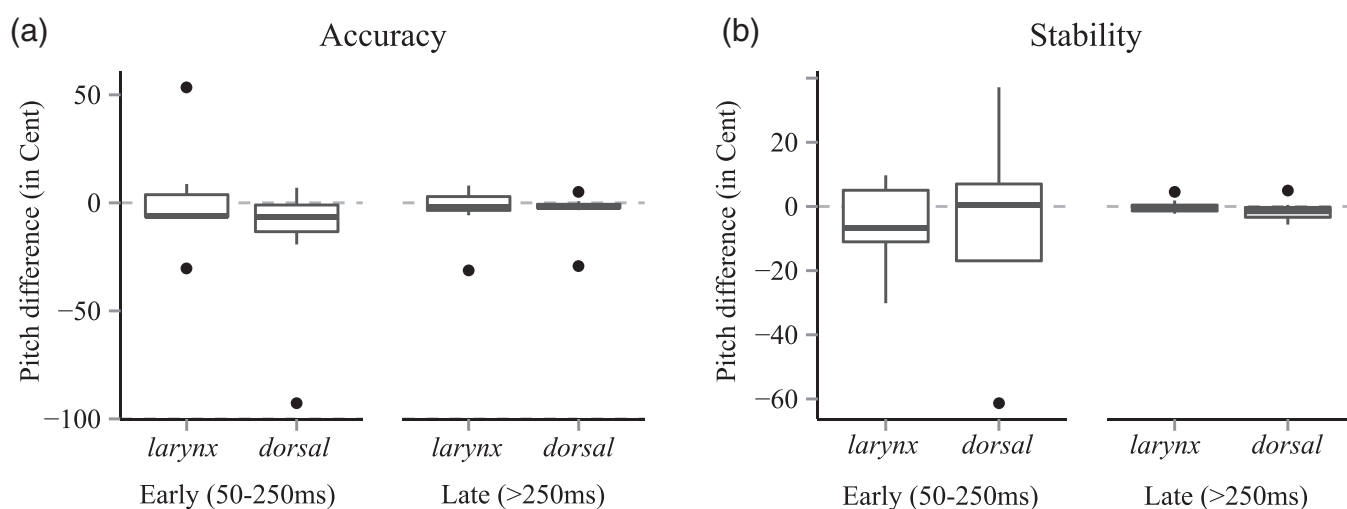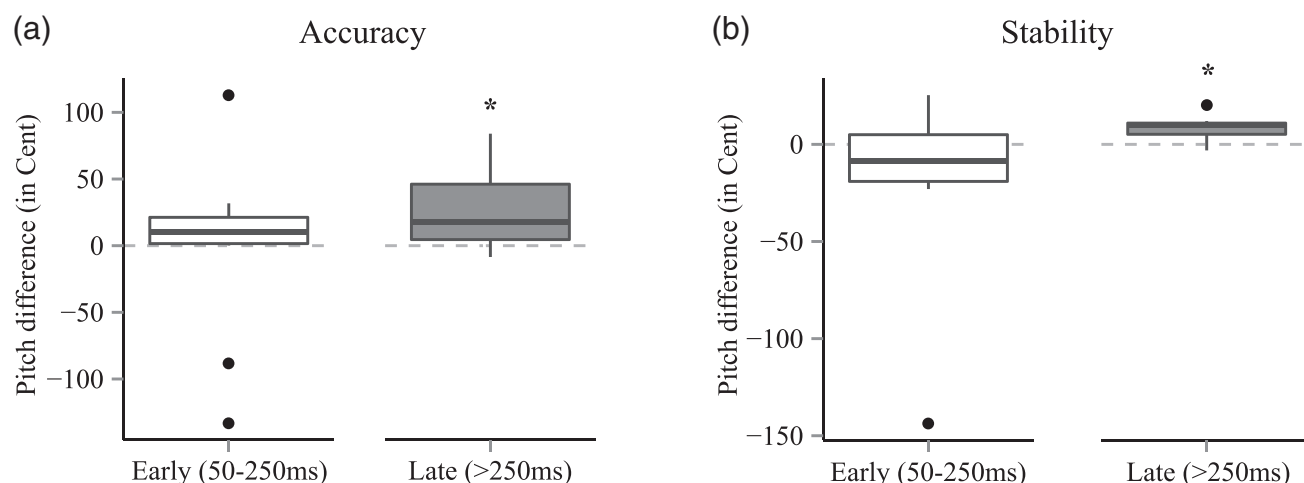

FIGURE 3 Results from Experiment I, comparing pitch-matching performance before and after iTBS over larynx-S1 and a dorsal-S1 control region (representing the lower limb). Participants could use auditory feedback to regulate vocal pitch. No significant effects were found, neither for (a) pitch accuracy nor for (b) pitch stability in the early (50–250 ms) and late (>250 ms) components after sound onset. Negative values indicate improvements (100 Cent = one semitone)

Exp. II: Effects of masked auditory feedback on pitch performance

(a) Accuracy

(b) Stability

Exp. II: Post *vs* pre iTBS changes in pitch performance (with masked hearing)
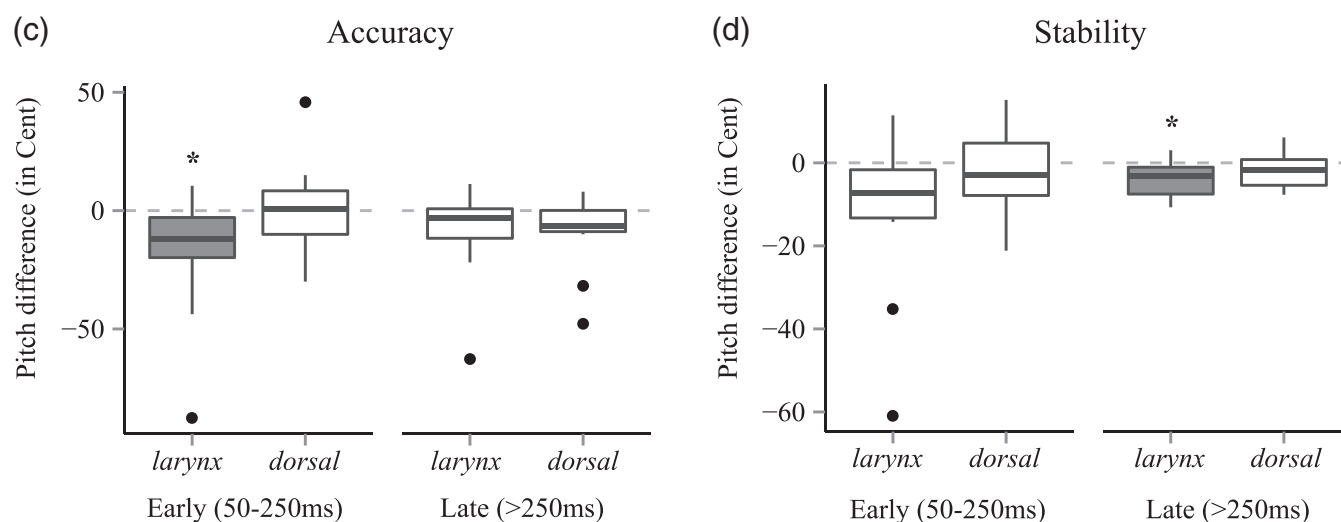
(c) Accuracy

(d) Stability

FIGURE 4 Results from Experiment II, comparing pitch-matching performance before and after iTBS over larynx-S1 and dorsal-S1 (representing the lower limb). Auditory self-monitoring was masked with loud pink noise during singing. Compared to singing with normal auditory feedback (prior to iTBS), masking significantly reduced (a) pitch accuracy and (b) pitch stability (>250 ms after sound onset). When auditory feedback was masked, larynx-S1 iTBS significantly improved (c) pitch accuracy (50–250 ms after sound onset) and (d) stability (>250 ms after sound onset). Significance is indicated by darker shadings. Significances: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

were calculated to test if either pitch-discrimination or pitch-matching accuracy levels prior to iTBS could moderate the behavioral effects of iTBS on pitch-matching performance.

3 | RESULTS

3.1 | Pitch-discrimination accuracy

All participants showed average to very-good pitch discrimination accuracy. In Experiment I, the perceived frequency difference was on average 0.5% (8.6 Cent), ranging between 0.17 and 1.08% (2.9–18.6 Cent). In Experiment II, it was 0.7% (12.8 Cent), ranging between 0.2 and 1.25% (3.5–21.5 Cent). As the smallest musical interval in the pitch-matching task was 100 Cent (1 semitone), this test ruled out

that poor pitch-discrimination accuracy could account for poor pitch-matching performance.

3.2 | Effects of iTBS on pitch-matching performance

In Experiment I, participants could monitor auditory feedback from their own voice for controlling pitch levels. As pitch-matching performance did not differ significantly between pre-iTBS sessions (i.e., prelarynx and predorsal S1 stimulation), results were averaged to facilitate statistical evaluation of post-iTBS related effects and reduce data dimensionality. Friedman's ANOVA revealed no significant effect of iTBS stimulation on pitch-matching accuracy (Figure 3a), neither for the initial [χ^2 (3, $N = 8$) = 5.10, $p = 0.17$] nor for the late time-window after sound onset [χ^2 (3, $N = 8$) = 1.65, $p = 0.65$]. As depicted in Figure 3b, iTBS stimulation also showed no significant effect on pitch stability,

TABLE 1 Results of the Wilcoxon sign rank test (one-sided, FDR corrected) in Experiment II, comparing pitch-matching performance while singing with and without masked auditory feedback from the own voice. * $p > 0.5$

Effects of masking-noise on pitch performance			
		Mask vs. no-mask	
Variable		z-value	p value
Early	Pitch accuracy	1.156	0.17
Late		2.401	**0.02***
Early	Pitch stability	−0.889	0.81
Late		2.756	**0.01***

neither for the initial [χ^2 (3, $N = 8$) = 2.40, $p = 0.49$] nor for the late time-window after sound onset [χ^2 (3, $N = 8$) = 1.05, $p = 0.79$.].

In Experiment II, auditory feedback from the own voice was masked with pink noise to reveal effects of iTBS on pitch-matching performance. Compared to singing with normal auditory feedback prior to iTBS in Experiment II, noise-masking significantly impaired pitch accuracy and pitch stability (>250 ms after sound onset), indicating that acoustic self-monitoring was substantially reduced during singing by the masking noise (Figure 4a,b). This effect was comparable across sessions (i.e., prior to larynx and dorsal S1 stimulation). Therefore, pre-iTBS results with masked singing were averaged to facilitate statistical evaluation of post-iTBS related effects and reduce data dimensionality. Table 1 depicts the results of the Wilcoxon sign-rank tests.

In Experiment II (Figure 4c,d), Friedman's ANOVA revealed a significant main effect of iTBS on pitch-matching accuracy for the early (50–250 ms; χ^2 [2, $N = 12$] = 6.1667, $p < 0.05*$) but not the late time-window after sound onset (>250 ms; χ^2 [2, $N = 12$] = 3.8723, $p = 0.14$). Conversely, iTBS showed a significant main effect on pitch stability for the late (>250 ms; χ^2 [2, $N = 13$] = 7.1667, $p < 0.05*$) but not the early time-window after sound onset (50–250 ms; χ^2 [2, $N = 12$] = 5.1667, $p = 0.08$). Post hoc comparisons (Table 2) revealed that these effects were due to right larynx-S1 iTBS. These result were underlined by a medium effect size for both pitch accuracy and pitch stability ($r = 0.48$ and $r = 0.51$; see Table 2).

3.3 | Exploratory correlation analyses

Correlation analyses in Experiment II suggest that participants with poorer pitch-matching accuracy during noise-masked singing prior to iTBS showed greater improvements after larynx-S1 iTBS for early ($r = −0.58$, $p < 0.05*$) and late ($r = −0.66$, $p < 0.01**$) pitch accuracy (Figure 5a,b).

Moreover, participants with larger (i.e., less accurate) pitch-discrimination thresholds showed stronger effects of larynx-S1 iTBS on early [$r = −0.73$, $p < 0.01**$] and late [$r = −0.66$, $p < 0.05*$] pitch-matching accuracy (Figure 4c,d) but were also more negatively affected by masking noise (Figure 5e) on late pitch-matching accuracy (>250 ms; $r = 0.85$, $p < 0.001***$) .
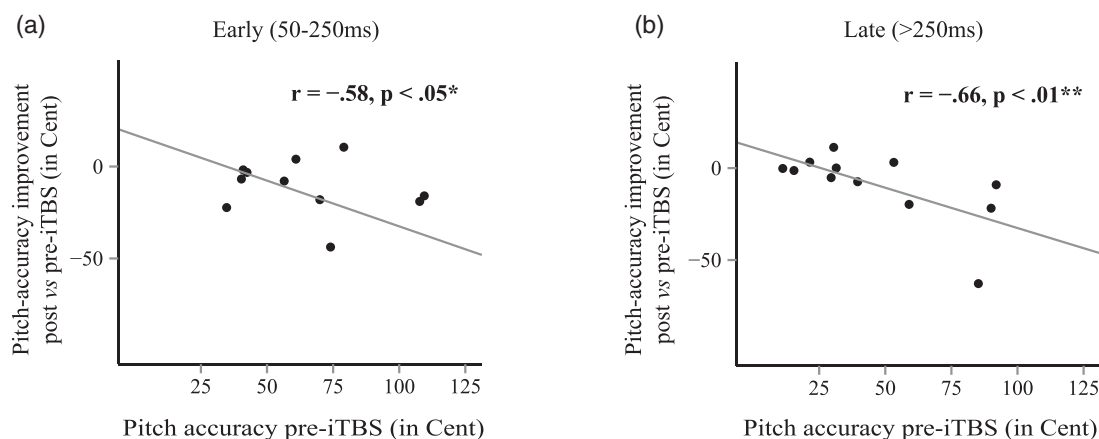
4 | DISCUSSION

We applied neuronavigated iTBS in musically untrained participants and found evidence for an involvement of right larynx S1 in pitch regulation during singing. Particularly, we observed enhanced pitch accuracy within an early (50–250 ms after sound onset) and enhanced pitch stability within a late (>250 ms after sound onset) time window after sound onset following right larynx S1 iTBS but not after stimulating a nonvocal S1 control area. Notably, this effect only occurred when acoustic feedback from the own voice was masked with noise.

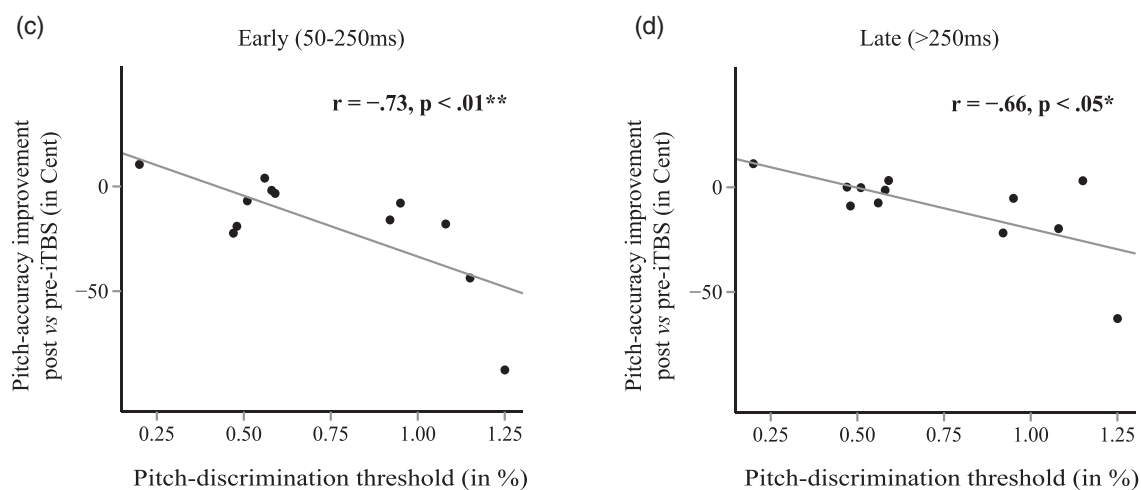## 4.1 | Temporal aspects of vocal pitch regulation

Studies have described a temporal pattern in vocal pitch regulation based on the observation that unexpected pitch shifts in acoustic feedback generate two discrete motor responses in the direction opposite to the shift, which reflect involuntary (~50–250 ms) and voluntary (≥250 ms) mechanisms of vocal motor control (Burnett et al., 1998; Burnett & Larson, 2002; Grell et al., 2009; Hain et al., 2000; Natke et al., 2003; Sapir, McClean, & Larson, 1983). Moreover, vocal intentions prior to task performance alter the voluntary but not the involuntary response characteristic (Hain et al., 2000), indicating that responses at voice onset are automatically controlled by efference mechanisms, whereas the later volitional responses compare current feedback to previous sensory production (Hawco & Jones, 2009).

Vocal compensatory effects have also been reported for pitch adjustments following mechanical displacements of the larynx (Loucks, Poletto, Saxon, & Ludlow, 2005; Sapir, Baker, Larson, & Ramig, 2000), revealing that kinesthetic error is processed faster and is also heavier weighted than auditory error during the early part of voice production (Larson et al., 2008). Vocalization onset is furthermore strongly determined by presetting laryngeal tension and respiratory pressure to specific values before audible sound is produced (Loucks et al., 2005; Wyke, 1974a). Given that larynx S1 iTBS enhanced pitch matching accuracy only at the onset of phonation, it is therefore feasible that the facilitation of vocalization

TABLE 2 Post hoc comparisons in experiment II, representing significant effects of iTBS over larynx-S1 relative to dorsal-S1 (i.e., lower limb) on early pitch-matching accuracy (50–250 ms after voice onset) and late pitch-matching stability (>250 ms after voice onset) when auditory self-monitoring was masked with noise. * $p > 0.5$

	Early pitch accuracy			Late pitch stability		
	Z	p	r	Z	p	r
Larynx-S1 iTBS (post vs. pre)	−2.35	**0.028***	0.48	−2.51	**0.028***	0.51
Dorsal-S1 iTBS (post vs. pre)	−0.08	0.479	–	−0.08	0.091	–
Larynx- vs. dorsal-S1 iTBS (post)	−2.00	**0.034***	0.41	−2.04	**0.031***	0.42

Effect sizes have been calculated using this formula: $r = Z/\sqrt{N}$.

**Correlations between pre-iTBS pitch-matching accuracy
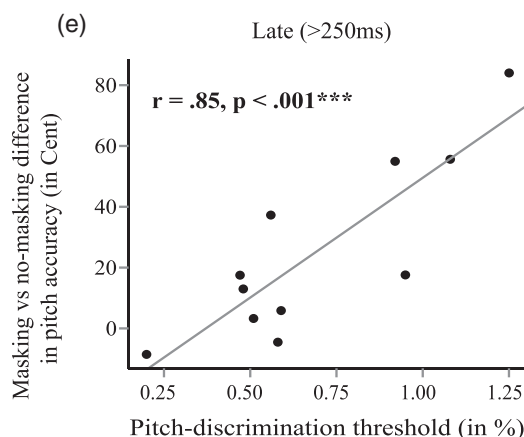and its improvement after larynx-S1 iTBS**

FIGURE 5 Correlation analyses from experiment II indicate that participants with poorer pitch-matching accuracy prior to iTBS showed greater improvements in pitch accuracy after larynx-S1 stimulation for (a) early (50–250 ms) and (b) late (>250 ms after sound onset) components after sound onset. Participants with larger pitch-discrimination thresholds (i.e., less accurate pitch discrimination) showed more after improvement after larynx-S1 iTBS in both (c) early and (d) late pitch-matching accuracy (post- pre-iTBS difference). They also showed stronger effects of (e) noise masking on pitch-matching accuracy (>250 ms after sound onset). Significances: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

related somatosensory processes supported efference mechanisms at utterance onset (Hawco & Jones, 2009) by integrating motor commands with proprioceptive and tactile information prior to vocalization (Bouchard et al., 2013).

4.2 | Role of sensorimotor interactions

That the effect of iTBS over the right laryngeal S1 area on vocal pitch regulation only occurs when auditory feedback was masked with pink noise can be accounted for by the interaction between auditory and somatosensory feedback. The neural processes for hearing perception undergo significant modulation by the somatosensation system (Wu, Stefanescu, Martel, & Shore, 2015) and vice versa (Ito & Ostry, 2012). Speakers are moreover sensitive to perturbations in both feedback subsystems, responding also to mechanical alterations of articulatory motion trajectories in the absence of auditory consequences (Lametti et al., 2012; Nasir & Ostry, 2006; Tremblay et al., 2003). This can be explained by the tight correlation between the simultaneously generated auditory and somatosensory feedback during speech production (Guenther & Vladusich, 2012; Ito et al., 2009; Ito & Ostry, 2012; Skipper et al., 2017). Studies have shown that the speech motor system may shift the weighting of the two sources of feedback depending on their reliability, to constrain the perturbation response and decrease the compensation (Katseff et al., 2012; Larson et al., 2008). The lack of auditory feedback may thus have triggered the monitoring system to attend more to somatosensory feedback, thereby facilitating the effect of larynx-S1 stimulation on pitch accuracy in the presence of pink-noise masking. However, it has also been suggested that M1 and S1 act as a low-level controller of articulatory movements, which in turn is controlled by a higher level SFC-based speech motor control network integrating both somatosensory and auditory feedback (Houde & Chang, 2015). Low-level interactions between M1 and S1 may thus explain why S1 stimulation only improved involuntary aspects of pitch regulation at sound onset but not at mid-utternace (Hawco & Jones, 2009). In contrast, rTMS over higher-order inferior parietal lobe altered also voluntary adaptive responses to feedback perturbations (Shum et al., 2011). An alternative explanation is that motor-to-sensory mappings in the context of musical pitch must first be accurately tuned through experience before somatosensory feedback may be used for higher-order volitional pitch regulation (Guenther, 2006; Tourville & Guenther, 2011).

4.3 | Role of experience

The observation that deaf individuals can produce intelligible speech in the complete absence of auditory feedback (Nasir & Ostry, 2008) indicates that the role of somatosensory feedback in speech motor coordination may depend on experience (Lametti et al., 2012). The DIVA model (Guenther & Vladusich, 2012) suggests that tactile-proprioceptive signals gain importance as a prime system for error detection after movement patterns have been sufficiently tuned by the auditory system. According to Friston et al. (Friston, 2011; Shipp, Adams, & Friston, 2013), somatosensory predictions represent the anticipated proprioceptive and kinesthetic consequences of the movement command issued by the agranular primary motor cortex (M1).

Within this framework, S1 generates probabilistic predictions about causes of sensory input that are continuously updated by prediction errors and relayed back to M1 to correct the descending motor commands (see also Barrett & Simmons, 2015). State-feedback control of pitch-goals based on somatosensory feedback may therefore be more accurate in singers, who show not only consistently higher pitch-matching skills when auditory feedback is available (Hutchins & Peretz, 2011, 2012; Nikjeh, Lister, & Frisch, 2009) but also sing more accurately than nonsingers both at the onset and during phonation when auditory feedback is masked (Murry, 1990; Watts, Murphy, & Barnes-Burroughs, 2003). Likewise, singers may also generate more accurate forward prediction of the somatosensory consequences of the motor command (Hickok et al., 2011; Perkell, 2012). Neuroimaging provides evidence for this notion, demonstrating that singers are not only less sensitive to auditory perturbations (Zarate & Zatorre, 2008) than to somatosensory perturbations (Kleber et al., 2013), but also that they are able to engage brain regions required for somatosensory feedback integration when auditory feedback is masked (Kleber et al., 2017). Conversely, nonsingers in that study were unable to compensate with somatosensory feedback and showed a disruption of brain regions involved in predictive/corrective mechanisms (Hickok, 2017; Houde & Chang, 2015). We therefore speculate that the contribution of S1 functions to higher-level state-feedback control may depend on experience, which could explain the limited effects of larynx-S1 stimulation on voluntary final pitch-regulation in nonsingers in Experiment II.

4.4 | Effects on pitch stability

Higher pitch stability during phonation indicates more even vocal fold oscillations (Titze, 1988; Titze, 2008), which is not only relevant for singing but also important for speech production (Orlikoff, 1995). Pitch stability stabilizes gradually throughout adolescence to adulthood (Boltezar, Burger, & Zargi, 1997; Smith, 2006). Importantly, rhythmically repetitive movements like sustained oscillatory patterns during phonation are driven by central pattern generator networks in the brainstem (Hage & Jurgens, 2006; Schoneich & Hedwig, 2012) and supported by somatosensory signals from mechanoreceptors in the vocal folds (Gozaine & Clark, 2005; Hammer & Barlow, 2010; Simonyan & Jurgens, 2005; Titze & Hunter, 2004; Wyke, 1974a, 1974b). Afferent information is relayed via brainstem and thalamic projections to the insular and somatosensory cortices and projected back via the solitary tract and the parabrachial nucleus onto medullary respiratory rhythm centers and phonatory motor neurons of the nucleus ambiguous and nucleus retroambiguus (Ackermann & Riecker, 2004; Eickhoff, Heim, Zilles, & Amunts, 2009; Jürgens, 2002). Elimination of these afferents leads to abnormal phonatory patterns (Shiba, Miura, Yuza, Sakamoto, & Nakajima, 1999; Shiba, Yoshida, & Miura, 1995). Given the importance of stable phonation in the context of speech motor learning, we speculate that a shift toward somatosensory feedback monitoring during masking (Katseff et al., 2012; Larson et al., 2008) allowed larynx-S1 stimulation to enhance state-feedback control strategies at mid-utterance (Hawco & Jones, 2009) to affect the voluntary control over temporal regularities in glottal cycle duration, irrespective of pitch accuracy.

4.5 | Moderating variables (correlations)

We found that the effects of larynx iTBS on pitch-matching accuracy (voluntary and involuntary components) were more pronounced in participants with lower pitch-matching accuracy before iTBS. This is a plausible interaction, as participants with lower entry-level skills commonly benefit more from interventions aiming at improving them (Ladda et al., 2014). Likewise, participants with lower pitch-discrimination accuracy showed greater improvement after right larynx-S1 stimulation and were also more negatively affected by auditory-feedback masking with respect to their pitch-matching accuracy. This effect can be accounted for by interactions between somatosensory and auditory feedback subsystems. One explanation is that the speech motor system of individuals with lower auditory perceptual skills may be driven by a stronger weighting of somatosensory feedback (Katseff et al., 2012; Lametti et al., 2012; Larson et al., 2008).

4.6 | Limitations

This study has several limitations that need to be considered. Firstly, we cannot rule out that neuronavigated stimulation of larynx S1 also affected adjacent regions within M1, particularly in the light of extensive connections between larynx S1 with M1 and parietal areas (Borich et al., 2015; Bouchard et al., 2013). Similarly, confirmation of our implicit assumption that the area targeted under the coil is also the area affected would require the combination of TMS with functional magnetic resonance imaging (Andoh & Zatorre, 2012). Secondly, we compared the effects of larynx-S1 iTBS to stimulation of a nonvocal control area but not to a cortical region that is involved in vocal motor control (e.g., IPL). Therefore, a double-dissociation was not possible, which would have strengthened our results. Moreover, future studies may consider including a real sham condition (Thielscher et al., 2011). Another limitation is the lack of standard audiometric testing to exclude the possibility that participants with lower auditory sensitivity benefitted more from iTBS stimulation. However, the fact that all participants showed normal pitch-discrimination abilities should guard us well against hearing deficits. Furthermore, we cannot exclude the possibility that a ceiling effect in Experiment I, in which participants could use auditory feedback to monitor pitch accuracy, has prevented us from finding effects on pitch performance. Finally, despite acceptable effect sizes, the small sample size of this study presents a limiting factor for the generalization of our results, particularly in Experiment I.

4.7 | Conclusions

In conclusion, we have shown that iTBS applied over right larynx S1 can improve pitch-matching performance in nonsingers, indicating that enhanced kinesthetic-proprioceptive processes contribute to vocal pitch regulation. More importantly, the finding that S1 stimulation affected involuntary (pitch accuracy) and voluntary (pitch stability) aspects of motor control differently lends support to our hypothesis that the hierarchical organization of neural mechanisms by which sensory feedback subsystems are integrated and compared to predictions during speech motor control may differ based on the individual's sensorimotor experience (see also Hickok, 2014; Houde & Chang, 2015).

ORCID

Martin Lotze https://orcid.org/0000-0003-4519-4956
Boris Kleber https://orcid.org/0000-0003-2975-4343

REFERENCES

Ackermann, H., & Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language*, *89*, 320–328.

Ambalavanar, R., Tanaka, Y., Selbie, W. S., & Ludlow, C. L. (2004). Neuronal activation in the medulla oblongata during selective elicitation of the laryngeal adductor response. *Journal of Neurophysiology*, *92*, 2920–2932.

Andoh, J., & Zatorre, R. J. (2012). Mapping the after-effects of theta burst stimulation on the human auditory cortex with functional imaging. *Journal of Visualized Experiments*, e3985. https://www.jove.com/video/3985/mapping-after-effects-theta-burst-stimulation-on-human-auditory

Andreatta, R. D., Mann, E. A., Poletto, C. J., & Ludlow, C. L. (2002). Mucosal afferents mediate laryngeal adductor responses in the cat. *Journal of Applied Physiology (Bethesda, MD: 1985)*, *93*, 1622–1629.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews. Neuroscience*, *16*, 419–429.

Bauer, J. J., & Larson, C. R. (2003). Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: Improvements in methodology for the pitch-shifting technique. *The Journal of the Acoustical Society of America*, *114*, 1048–1054.

Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A., & Larson, C. R. (2014). Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain and Cognition*, *84*, 97–108.

Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A., & Larson, C. R. (2015). Functional role of delta and theta band oscillations for auditory feedback processing during vocal pitch motor control. *Frontiers in Neuroscience*, *9*, 109.

Behroozmand, R., Oya, H., Nourski, K. V., Kawasaki, H., Larson, C. R., Brugge, J. F., ... Greenlee, J. D. (2016). Neural correlates of vocal production and motor control in human Heschl's Gyrus. *The Journal of Neuroscience*, *36*, 2302–2315.

Behroozmand, R., Phillip, L., Johari, K., Bonilha, L., Rorden, C., Hickok, G., & Fridriksson, J. (2018). Sensorimotor impairment of speech auditory feedback processing in aphasia. *NeuroImage*, *165*, 102–111.

Behroozmand, R., & Sangtian, S. (2018). Neural bases of sensorimotor adaptation in the vocal motor system. *Experimental Brain Research*, *236*, 1881–1895.

Belyk, M., & Brown, S. (2017). The origins of the vocal brain in humans. *Neuroscience and Biobehavioral Reviews*, *77*, 177–193.

Bennemann, J., Freigang, C., Schroger, E., Rubsamen, R., & Richter, N. (2013). Resolution of lateral acoustic space assessed by electroencephalography and psychoacoustics. *Frontiers in Psychology*, *4*, 338.

Berry, D. A., Montequin, D. W., Chan, R. W., Titze, I. R., & Hoffman, H. T. (2003). An investigation of cricoarytenoid joint mechanics using simulated muscle forces. *Journal of Voice*, *17*, 47–62.

Boltezar, I. H., Burger, Z. R., & Zargi, M. (1997). Instability of voice in adolescence: Pathologic condition or normal developmental variation? *The Journal of Pediatrics*, *130*, 185–190.

Borich, M. R., Brodie, S. M., Gray, W. A., Ionta, S., & Boyd, L. A. (2015). Understanding the role of the primary somatosensory cortex: Opportunities for rehabilitation. *Neuropsychologia*, 79, 246–255.

Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495, 327–332.

Breshears, J. D., Molinaro, A. M., & Chang, E. F. (2015). A probabilistic map of the human ventral sensorimotor cortex using electrical stimulation. *Journal of Neurosurgery*, 123, 340–349.

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2009). The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain and Cognition*, 70, 31–41.

Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18, 837–845.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, 103, 3153–3161.

Burnett, T. A., & Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America*, 112, 1058–1063.

Burnett, T. A., McCurdy, K. E., & Bright, J. C. (2008). Reflexive and volitional voice fundamental frequency responses to an anticipated feedback pitch error. *Experimental Brain Research*, 191, 341–351.

de Cheveigne, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111, 1917–1930.

Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 367, 2399–2421.

Friberg, A., Schoonderwaldt, E., & Juslin, P. N. (2005). CUEX: An algorithm for extracting expressive tone variables from audio recordings. *Acoustica United with Acta Acoustica*, 93, 411–420.

Friston, K. (2011). What is optimal about motor control? *Neuron*, 72, 488–498.

Golfinopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *NeuroImage*, 55, 1324–1338.

Gozaine, T. C., & Clark, K. F. (2005). Function of the laryngeal mechanoreceptors during vocalization. *Laryngoscope*, 115, 81–88.

Grabski, K., Lamalle, L., Vilain, C., Schwartz, J. L., Vallee, N., Tropres, I., ... Sato, M. (2012). Functional MRI assessment of orofacial articulators: Neural correlates of lip, jaw, larynx, and tongue movements. *Human Brain Mapping*, 33, 2306–2321.

Grell, A., Sundberg, J., Ternstrom, S., Ptok, M., & Altenmüller, E. (2009). Rapid pitch correction in choir singers. *The Journal of the Acoustical Society of America*, 126, 407–413.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39, 350–365.

Guenther, F. H. (2016). *Neural control of speech*. Cambridge, MA: The MIT Press xiv, 410 pages.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, 25, 408–422.

Hage, S. R., & Jurgens, U. (2006). On the role of the pontine brainstem in vocal pattern generation: A telemetric single-unit recording study in the squirrel monkey. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 26, 7105–7115.

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research*, 130, 133–141.

Hammer, M. J., & Barlow, S. M. (2010). Laryngeal somatosensory deficits in Parkinson's disease: Implications for speech respiratory and phonatory control. *Experimental Brain Research*, 201, 401–409.

Hawco, C. S., & Jones, J. A. (2009). Control of vocalization at utterance onset and mid-utterance: Different mechanisms for different goals. *Brain Research*, 1276, 131–139.

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language & Cognitive Processes*, 29, 2–20.

Hickok, G. (2017). A cortical circuit for voluntary laryngeal control: Implications for the evolution language. *Psychonomic Bulletin & Review*, 24, 56–63.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, 69, 407–422.

Houde, J. F., & Chang, E. F. (2015). The cortical computations underlying feedback control in vocal production. *Current Opinion in Neurobiology*, 33, 174–181.

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, 82.

Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron*, 45, 201–206.

Hutchins, S., & Peretz, I. (2011). Perception and action in singing. *Progress in Brain Research*, 191, 103–118.

Hutchins, S. M., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology. General*, 141, 76–97.

Ito, T., Coppola, J. H., & Ostry, D. J. (2016). Speech motor learning changes the neural response to both auditory and somatosensory signals. *Scientific Reports*, 6, 25926.

Ito, T., & Ostry, D. J. (2012). Speech sounds alter facial skin sensation. *Journal of Neurophysiology*, 107, 442–447.

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 1245–1248.

Jacobs, M., Premji, A., & Nelson, A. J. (2012). Plasticity-inducing TMS protocols to investigate somatosensory control of hand function. *Neural Plasticity*, 2012, 350574.

Jones, C. B., Lulic, T., Bailey, A. Z., Mackenzie, T. N., Mi, Y. Q., Tommerdahl, M., & Nelson, A. J. (2016). Metaplasticity in human primary somatosensory cortex: Effects on physiology and tactile perception. *Journal of Neurophysiology*, 115, 2681–2691.

Jones, J. A., & Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Experimental Brain Research*, 190, 279–287.

Jurgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26, 235–258.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26, 235–258.

Kalla, R., Muggleton, N. G., Cowey, A., & Walsh, V. (2009). Human dorsolateral prefrontal cortex is involved in visual search for conjunctions but not features: A theta TMS study. *Cortex*, 45, 1085–1090.

Katayama, T., & Rothwell, J. C. (2007). Modulation of somatosensory evoked potentials using transcranial magnetic intermittent theta burst stimulation. *Clinical Neurophysiology*, 118, 2506–2511.

Katseff, S., Houde, J., & Johnson, K. (2012). Partial compensation for altered auditory feedback: A tradeoff with somatosensory feedback? *Language and Speech*, 55, 295–308.

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *NeuroImage*, 36, 889–900.

Kleber, B., Friberg, A., Zeitouni, A., & Zatorre, R. (2017). Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-masked auditory feedback in singers and nonsingers. *NeuroImage*, 147, 97–110.

Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., & Lotze, M. (2010). The brain of opera singers: Experience-dependent changes in functional activation. *Cerebral Cortex*, 20, 1144–1152.

Kleber, B., Veit, R., Moll, C. V., Gaser, C., Birbaumer, N., & Lotze, M. (2016). Voxel-based morphometry in opera singers: Increased gray-matter volume in right somatosensory and auditory cortices. *NeuroImage*, 133, 477–483.

Kleber, B., Zeitouni, A. G., Friberg, A., & Zatorre, R. J. (2013). Experience-dependent modulation of feedback integration during singing: Role of the right anterior insula. *The Journal of Neuroscience*, 33, 6070–6080.

Kort, N. S., Cuesta, P., Houde, J. F., & Nagarajan, S. S. (2016). Bihemispheric network dynamics coordinating vocal feedback control. *Human Brain Mapping*, 37, 1474–1485.

Kuypers, H. G. (1958a). An anatomical analysis of cortico-bulbar connexions to the pons and lower brain stem in the cat. *Journal of Anatomy*, 92, 198–218.

Kuypers, H. G. (1958b). Corticobular connexions to the pons and lower brain-stem in man: An anatomical study. *Brain*, *81*, 364–388.

Kuypers, H. G. (1958c). Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *The Journal of Comparative Neurology*, *110*, 221–255.

Ladda, A. M., Pfannmoeller, J. P., Kalisch, T., Roschka, S., Platz, T., Dinse, H. R., & Lotze, M. (2014). Effects of combining 2 weeks of passive sensory stimulation with active hand motor training in healthy adults. *PLoS One*, *9*, e84402.

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *The Journal of Neuroscience*, *32*, 9351–9358.

Larson, C. R., Altman, K. W., Liu, H., & Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice F0 control. *Experimental Brain Research*, *187*, 613–621.

Lee, K. G., Jacobs, M. F., Asmussen, M. J., Zapallow, C. M., Tommerdahl, M., & Nelson, A. J. (2013). Continuous theta-burst stimulation modulates tactile synchronization. *BMC Neuroscience*, *14*, 89.

Leveque, Y., Muggleton, N., Stewart, L., & Schon, D. (2013). Involvement of the larynx motor area in singing-voice perception: A TMS study(dagger). *Frontiers in Psychology*, *4*, 418.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*(Suppl 2), 467+.

Leydon, C., Bauer, J. J., & Larson, C. R. (2003). The role of auditory feedback in sustaining vocal vibrato. *The Journal of the Acoustical Society of America*, *114*, 1575–1581.

Liu, H., Behroozmand, R., Bove, M., & Larson, C. R. (2011). Laryngeal electromyographic responses to perturbations in voice pitch auditory feedback. *The Journal of the Acoustical Society of America*, *129*, 3946–3954.

Lotze, M., Seggewies, G., Erb, M., Grodd, W., & Birbaumer, N. (2000). The representation of articulation in the primary sensorimotor cortex. *Neuroreport*, *11*, 2985–2989.

Loucks, T. M., Poletto, C. J., Saxon, K. G., & Ludlow, C. L. (2005). Laryngeal muscle responses to mechanical displacement of the thyroid cartilage in humans. *Journal of Applied Physiology (Bethesda, MD: 1985)*, *99*, 922–930.

Ludlow, C. L. (2005). Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology & Neurobiology*, *147*, 205–222.

Masaki, H., Tanaka, H., Takasawa, N., & Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *Neuroreport*, *12*, 1851–1855.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*, 36–47.

Mor, N., Simonyan, K., & Blitzer, A. (2018). Central voice production and pathophysiology of spasmodic dysphonia. *Laryngoscope*, *128*, 177–183.

Morley, J. W., Vickery, R. M., Stuart, M., & Turman, A. B. (2007). Suppression of vibrotactile discrimination by transcranial magnetic stimulation of primary somatosensory cortex. *The European Journal of Neuroscience*, *26*, 1007–1010.

Mürbe, D., Pabst, F., Hofmann, G., & Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback: A longitudinal study of singers' pitch control. *Journal of Voice*, *18*, 236–241.

Murry, T. (1990). Pitch-matching accuracy in singers and nonsingers. *Journal of Voice*, *4*, 317–321.

Nasir, S. M., & Ostry, D. J. (2006). Somatosensory precision in speech production. *Current Biology*, *16*, 1918–1923.

Nasir, S. M., & Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature Neuroscience*, *11*, 1217–1222.

Natke, U., Donath, T. M., & Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *The Journal of the Acoustical Society of America*, *113*, 1587–1593.

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). The relationship between pitch discrimination and vocal production: Comparison of vocal and instrumental musicians. *The Journal of the Acoustical Society of America*, *125*, 328–338.

Orlikoff, R. F. (1995). Vocal stability and vocal tract configuration: An acoustic and electroglottographic investigation. *Journal of Voice*, *9*, 173–181.

Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *NeuroImage*, *33*, 628–635.

Parkinson, A. L., Behroozmand, R., Ibrahim, N., Korzyukov, O., Larson, C. R., & Robin, D. A. (2014). Effective connectivity associated with auditory error detection in musicians with absolute pitch. *Frontiers in Neuroscience*, *8*, 46.

Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., & Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage*, *61*, 314–322.

Parrell, B., Lammert, A. C., Ciccarelli, G., & Quatieri, T. F. (2017). Current speech motor control models: An overview of architectures & properties. *bioRxiv*. https://www.biorxiv.org/.

Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, *25*, 382–407.

Pfannmoller, J. P., Schweizer, R., & Lotze, M. (2016). Automated analysis protocol for high resolution BOLD-fMRI mapping of the fingertip somatotopy in brodmann area 3b. *Journal of Magnetic Resonance Imaging: JMRI*, *43*, 479–486.

Ragert, P., Franzkowiak, S., Schwenkreis, P., Tegenthoff, M., & Dinse, H. R. (2008). Improvement of tactile perception and enhancement of cortical excitability through intermittent theta burst rTMS over human primary somatosensory cortex. *Experimental Brain Research*, *184*, 1–11.

Rai, N., Premji, A., Tommerdahl, M., & Nelson, A. J. (2012). Continuous theta-burst rTMS over primary somatosensory cortex modulates tactile perception on the hand. *Clinical Neurophysiology*, *123*, 1226–1233.

Reichenbach, A., Thielscher, A., Peer, A., Bulthoff, H. H., & Bresciani, J. P. (2014). A key region in the human parietal cortex for processing proprioceptive hand feedback during reaching movements. *NeuroImage*, *84*, 615–625.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*, *11*, 1997–2000.

Rohl, M., Kollmeier, B., & Uppenkamp, S. (2011). Spectral loudness summation takes place in the primary auditory cortex. *Human Brain Mapping*, *32*, 1483–1496.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York, NY: Russell Sage Foundation.

Roux, F. E., Djidjeli, I., & Durand, J. B. (2018). Functional architecture of the somatosensory homunculus detected by electrostimulation. *The Journal of Physiology*, *596*, 941–956.

Sapir, S., Baker, K. K., Larson, C. R., & Ramig, L. O. (2000). Short-latency changes in voice F0 and neck surface EMG induced by mechanical perturbations of the larynx during sustained vowel phonation. *Journal of Speech, Language, and Hearing Research*, *43*, 268–276.

Sapir, S., McClean, M. D., & Larson, C. R. (1983). Human laryngeal responses to auditory stimulation. *The Journal of the Acoustical Society of America*, *73*, 315–321.

Schoneich, S., & Hedwig, B. (2012). Cellular basis for singing motor pattern generation in the field cricket (Gryllus bimaculatus DeGeer). *Brain and Behavior: A Cognitive Neuroscience Perspective*, *2*, 707–725.

Shiba, K., Miura, T., Yuza, J., Sakamoto, T., & Nakajima, Y. (1999). Laryngeal afferent inputs during vocalization in the cat. *Neuroreport*, *10*, 987–991.

Shiba, K., Yoshida, K., & Miura, T. (1995). Functional roles of the superior laryngeal nerve afferents in electrically induced vocalization in anesthetized cats. *Neuroscience Research*, *22*, 23–30.

Shipp, S., Adams, R. A., & Friston, K. J. (2013). Reflections on agranular architecture: Predictive coding in the motor cortex. *Trends in Neurosciences*, *36*, 706–716.

Shum, M., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2011). Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *The European Journal of Neuroscience*, *34*, 1817–1822.

Simon, R. (1989). Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials*, *10*, 1–10.

Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *The Neuroscientist*, *17*, 197–208.

Simonyan, K., & Jurgens, U. (2005). Afferent cortical connections of the motor cortical larynx area in the rhesus monkey. *Neuroscience*, *130*, 133–149.

Skillings, J. H., & Mack, G. A. (1981). On the use of a Friedman-type statistic in balanced and unbalanced block designs. *Technometrics*, *23*(2), 171–177.

Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain and Language, 164,* 77–105.

Smith, A. (2006). Speech motor development: Integrating muscles, movements, and linguistic units. *Journal of Communication Disorders, 39,* 331–349.

Smotherman, M. S. (2007). Sensory feedback control of mammalian vocalizations. *Behavioural Brain Research, 182,* 315–326.

Sundberg, J., Iwarsson, J., & Billstrom, A. H. (1995). Significance of mechanoreceptors in the subglottal mucosa for subglottal pressure control in singers. *Journal of Voice, 9,* 20–26.

Thielscher, A., Opitz, A., & Windhoff, M. (2011). Impact of the gyral geometry on the electric field induced by transcranial magnetic stimulation. *NeuroImage, 54,* 234–243.

Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *The Journal of the Acoustical Society of America, 83,* 1536–1552.

Titze, I. R. (2008). Nonlinear source-filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America, 123,* 2733–2749.

Titze, I. R., & Hunter, E. J. (2004). Normal vibration frequencies of the vocal ligament. *The Journal of the Acoustical Society of America, 115,* 2264–2269.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language & Cognitive Processes, 26,* 952–981.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39,* 1429–1443.

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature, 423,* 866–869.

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America, 122,* 2306–2319.

Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: Report and suggested guidelines from the international workshop on the safety of repetitive transcranial magnetic stimulation, June may 7, 1996. *Electroencephalography and Clinical Neurophysiology, 108,* 1–16.

Watts, C., Murphy, J., & Barnes-Burroughs, K. (2003). Pitch matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with nontalented singing voices in conditions of varying feedback. *Journal of Voice, 17,* 185–194.

Wu, C., Stefanescu, R. A., Martel, D. T., & Shore, S. E. (2015). Listening to another sense: Somatosensory integration in the auditory system. *Cell and Tissue Research, 361,* 233–250.

Wyke, B. D. (1974a). Laryngeal neuromuscular control systems in singing. A review of current concepts. *Folia Phoniatrica (Basel), 26,* 295–306.

Wyke, B. D. (1974b). Proceedings: Laryngeal myotatic reflexes and phonation. *Folia Phoniatrica (Basel), 26,* 249–264.

Yoshida, Y., Mitsumasu, T., Hirano, M., & Kanaseki, T. (1985). Somatotopic representation of the laryngeal motoneurons in the medulla of monkeys. *Acta Oto-Laryngologica, 100,* 299–303.

Zarate, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience, 7,* 237.

Zarate, J. M., Delhommeau, K., Wood, S., & Zatorre, R. J. (2010). Vocal accuracy and neural plasticity following micromelody-discrimination training. *PLoS One, 5,* e11181.

Zarate, J. M., Wood, S., & Zatorre, R. J. (2010). Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia, 48,* 607–618.

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage, 40,* 1871–1887.

Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology, 10,* e1001372.